

Weekly Report

Sept. 24th, 2018

Done:

1. 继续上周的草图设计, 结合前两周阅读的两篇 design space 综述. 人在视图寻找一个通顺的逻辑来概括这些设计(总结一个言之有理的套路), 但是还没有找到. 等马老师回来再问问是否有一些理解上的不一致. 其他一些事宜:

在 Google slides 上画了一些草图.

原本一些模棱两可的(也就是既可以 data fusion 也可以 visual fusion 的)草图被纠正.

两者关系紧密, 我觉得两者还是要结合, 也就是 DF, VF 和 DF&VF 三者都要存在.

尽管只是 visual report, 仍然需要一定工作量.

2. 按照官方教程, 复习 tensorflow keras 的基本操作, 给空调组的小朋友们分配了一些数据处理和机器学习入门相关的任务.

To Do:

1. 在 DF&VF 方面, 之前没考虑但又不可避免, 需要一些东西来说明为什么要有或者为什么没有

2 从 user study 角度而言, 我们现在明显的看到了有些东西只能用 DF 有些只能 VF, 有些两者都能用, 有些则模棱两可, 那对此采取的测试也就不一样. 要采取什么测试, 找了梅博写的套路, 但好像仍不太顺, 需要进一步调研.

Paper Reading:

复习了 k-means 聚类, 其最后解也是一个优化问题.

在此基础上粗略研究了一下谱聚类的文章, 他比较实用但是比较复杂还没有完全看懂.

KDD18 *Discovering Non-Redundant K-means Clusterings in Optimal Subspaces* 简单的说是一种在不同的特征子空间里做聚类的办法. 比如给你一些物体, 在形状这个维度上可以有一种聚类, 在颜色上可以有另一种聚类. 通过本文的办法能够一次性找出来并且能发现一个不好用来聚类的 noise subspace. 这个聚类最后也变成一个优化问题的解法.

CHI18 *Somewhere Over the Rainbow: An Empirical Assessment of Quantitative Colormaps* 对于 color map 的一个评测. Color map 一般用于量化的编码, 用于区分不同的值. 关于哪些 color map 更好用(速度, 准确率)等进行一个评估; 作者还发现如果加上 color naming 会提升一点效果.